2021 IEEE/ACIS 22nd International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD) | 978-1-6654-0403-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/SNPD51163.2021.9704884

Deep Learning Used to Recognition Swimmers Drowning

Jia-Xian Jian
National Chin-Yi University of Technology
Computer Science and Information Engineering
Taichung City, Taiwan
allensa119625@gmail.com

Chuin-Mu Wang
National Chin-Yi University of Technology
Computer Science and Information Engineering
Taichung City, Taiwan
cmwang@ncut.edu.tw

***Abstract*—Many people believe that when drowning occurs, there will be calls for help. In fact, people who are drowning do not get too many splashes or cry for help. They only try to get themselves out of the water by treading on the water. The drowning condition may cause serious brain damage, so it is extremely important to shorten the time it takes to detect the occurrence of drowning and rescue.**

This paper proposes using computer image processing technology to introduce artificial intelligence motion technology, mounting the camera on the bottom of the swimming pool, and use OpenPose to mark the image joint point features, and input the captured joint point features into the recursive neural network to determine whether the swimmer is drowning . The final training result is about 89.4% accurate, so it can be used to assist on-site lifeguards to detect swimmers who may be drowning, and to reduce incidents that cannot be detected immediately

Keywords—OpenPose、Thin-MobileNet、Drown recognition、Skeleton Based.

I. INTRODUCTION

Action recognition is mainly divided into three stages, namely the joint point labeling stage, the joint data extraction stage, and the neural network training stage. This article uses the OpenPose algorithm [1] in the joint skeleton labeling stage, which is different from the traditional OpenPose internal network model. Use VGG-19 to capture image features and replace the model with thin-MobileNet [2] to label the human skeleton joint points in the image. After the joint data is extracted in the second stage, the skeleton and joint features will become the third stage action recognition neural network. The input of the road, the third stage is to use shallow RNN network for training and identification.

Collected in the swimming pool drowning posture database, researched and analyzed the characteristics of drowning action on the Internet in advance, and went to the school swimming pool to simulate the possible posture of drowning, and collected training and test data required for action recognition.

II. TECHNICAL DESCRIPTION

When a drowning situation occurs, the drowning person will not be able to breathe or call for help because the mouth will be submerged in the water and then surfaced, so the call for help cannot be used to detect the occurrence of the drowning situation. This article uses the posture to determine whether the swimmer has drowned.

A. Drowning Pose

In the Brett & Kate McKay drowning survival guide[3], it is mentioned that swimmers will not have violent struggling movements when drowning occurs. The drowning person’s arm will press down on the water surface sideways to try to keep himself on the water surface. In addition, drowning It is common for people to crawl on an invisible ladder, which is often considered to be a normal kick in the pool.

In addition to studying the possible postures of drowning people when drowning, Rodman Louis on the Internet uploaded on the YouTube platform the results of the recognition of drowning movements by his sports AI laboratory [6]. The drowning movements of the research results have also become our subsequent drowning movements.

In this article, in addition to studying the text data on the Internet, this article also refers to the recognition results of other authors' drowning motion detection. It is hoped that the database of drowning motions will be as close as possible to the real drowning motion.

B. Skeleton Joint Point Annotation

In the joint skeleton icon annotation algorithm, OpenPose[1] and AlphaPose[4] are commonly used today. Although AlphaPose[4] is in the MPII human body motion database, the accuracy of different parts is higher than that of OpenPose[1] as shown in Table 1, but it consumes GPU memory resources. Or in terms of processing speed, OpenPose[1] is higher than AlphaPose[4]. As shown in Table 2, the time spent in finding the occurrence of drowning and the rescue is extremely important, so this article uses OpenPose[1] to mark the skeleton joint points.

TABLE I. COMPARISON OF OPENPOSE AND ALPHAPOSE ACCURACY [5]

Method	Head	Elbow	Wrist	Knee	Ankle
OpenPose	91.2	77.7	66.8	68.9	75.6
AlphaPose	91.3	84.0	76.4	79.9	82.1

TABLE II. OPENPOSE AND ALPHAPOSE PERFORMANCE COMPARISON [4]

Multiple person		
Method	GPU Memory Usage	Processing Speed
OpenPose	21.3%	18.39
AlphaPose	73.4%	1.15

Because OpenPose[1] belongs to the Bottom-up algorithm, the camera rack is set under the water surface and a simulation image of drowning is recorded. The image is marked by OpenPose[1]. As shown in Figure 1, it is found that when the drowning action occurs, the limbs are swinging and kicking in the water. Water will generate bubbles under the water sheet, and the bubbles in the water will obscure the skeleton position of the human body. Therefore, when using OpenPose[1] to mark the skeleton joint points, some pictures will not be marked. This is the reason why the drowning action is sometimes recognized incorrectly.

In addition, in order to improve the performance of OpenPose[1] skeleton joint annotation, the VGG-19 network used in OpenPose[1] is replaced by Thin-MobileNet[2] proposed in 2019. This network does not sacrifice accuracy on the premise of reducing the size of the network model. It is reduced to 9.9MB, which is much smaller than the 39.1MB model size of MobileNet-V1. The calculation time has also been reduced from the original 31 seconds to 14 seconds. After the change, it has sacrificed some marking accuracy but improved the marking performance of skeleton joint points.

Fig. 1. Water Splash Affects Skeleton Joint Annotation

In the annotation of skeleton joint points, although OpenPose[1] will be affected by splashes, the OpenPose[1] of the Thin-MobileNet[2] model in the validation set can still mark the skeleton joint points of most swimmers in the screen while improving the annotation performance, as shown in Figure 2. Show.

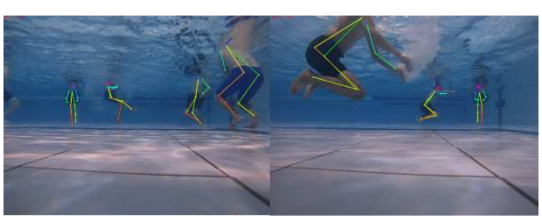

Fig. 2. Multi-Person Skeleton Joint Point Labeling

C. Feature Extraction of Skeleton Joint Points

After the previous stage OpenPose[1] marked the position of the skeleton joint points in the image, a total of 13 skeleton joint point positions will be marked. Next, the X-axis and Y-axis coordinates of each position of the skeleton joint points will be stored. The angular relationship between the limbs and the main structure of the human body will be calculated, and it will be stored in a one-dimensional array with the label of the image at the same time as the input data of the next stage of neural network.

D. Neural Network Training

Because GPU memory resources are already occupied during the stage of skeleton joint annotation, it is necessary to minimize the delay in drowning detection. Therefore, in the overall architecture, the skeleton joint point feature information extracted by OpenPose[1] is input into shallow In the layered RNN network, as shown in Figure 3, the training is performed, and the trained model is used for subsequent recognition of drowning actions.

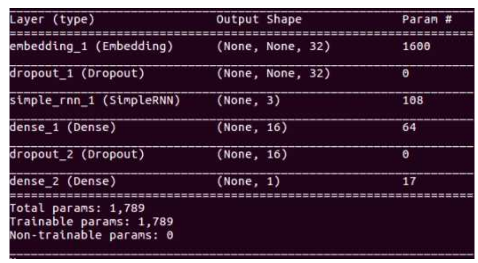

Fig. 3. Shallow RNN Network Architecture

E. Overall Recognition Framework

When the camera captures the drowning image, the image will be marked with the skeleton joint points in the Thin-MobileNet model's OpenPose algorithm. After marking the position of the skeleton joint points, the information such as the joint point coordinates will be taken as input to the shallow layer In the pre-trained model of the RNN network, it can recognize whether the characters in the screen have drowning actions. The recognition results are mainly divided into two types: drowning and normal, Use image recognition technology to assist on-site lifeguards to determine whether swimmers in the water are drowning, and reduce the occurrence of drowning that the lifeguards are too late to notice. The recognition flowchart is shown in Figure 4.

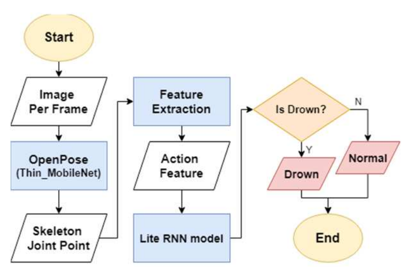

Fig. 4. Flow Chart of Drowning Recognition

III. Experimental Results

Under the action recognition in this paper, a simulation image containing drowning images was taken in the swimming pool for about 4 hours for drowning recognition training. There are lifeguards at the shooting scene, and the subject of the video has considerable swimming skills. Therefore, when shooting simulated images, the safety of swimmer can be guaranteed.The drowning action training data screen is shown in Figure 5.

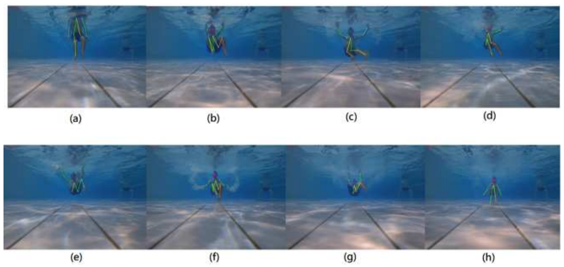

Fig. 5. Drowning Training Dataset Diagram

After the skeletal features are extracted, the image will be input into the shallow RNN network for training and a model for subsequent verification set testing will be obtained.

TABLE III. Confuse Matrix

Predicted / Actual	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (1a)$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (1b)$$

$$\text{F1-score} = \frac{2\times\text{Precision}\times\text{Recall}}{(\text{Precision}+\text{Recall})} \quad (1c)$$

After training through the network and substituting it into the test, the output confusion matrix is shown in Figure 6 below. After the computer output verification result, the final test result accuracy is 89.4%. The output Precision, Recall, F1-score, such as Table 4 shows.

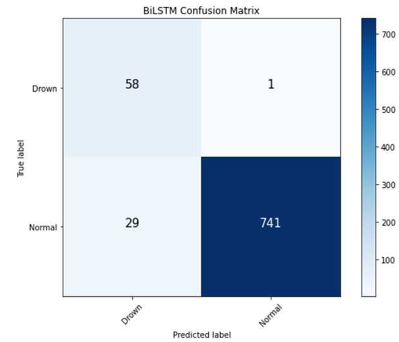

Fig. 6. Confusion Matrix Diagram

TABLE IV. Model Evaluation Results

	Precision	Recall	F1-Score
Drown	0.98	0.67	0.79
Normal	0.96	1.00	0.98
Macro AVG	0.97	0.83	0.89

After the test images are marked with skeleton joint points through OpenPose, and after the skeleton joint features are extracted and input into the pre-trained model, the swimmer's action in the screen can be judged as drowning or normal, as shown in Figure 7.

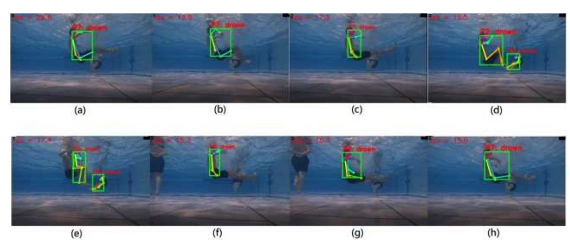

Fig. 7. Recognition Result Diagramm

IV. Conclusion

This paper applies current motion recognition technology to the recognition of drowning in swimming pools, hoping to prevent lifeguards from being unaware of the drowning incident. In the training set, the drowning simulation motion is mainly input to the OpenPose of the Thin-MobileNet[2] model to mark the skeleton joint points. Then, use the shallow RNN trained in advance to identify the neural network. The computer calculation results show that the accuracy of the drowning action recognition in this paper is 89.4%.

The actual test images found that if there are several continuous images in the general swimming or splashing action that are similar to the drowning action in the training set, it will be recognized as a drowning situation. In addition, too much air bubbles generated by the drowning swimmer in the water will also occur. There is a chance that the action cannot be captured by the computer

But in most situations, this paper simulates the possible actions of swimmers when a drowning event occurs, and uses the method proposed in this article to identify it. It is found that for simulated images, in most cases, it is possible to distinguish whether the drowning event occurred in the image.

REFERENCES

[1] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei and Yaser Sheikh "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," Computer Vision and Pattern Recognition, May 2019.

[2] Debjyoti Sinha and Mohamed El-Sharkawy, "Thin MobileNet: An Enhanced MobileNet Architecture," 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), Oct 2019.

[3] Brett and Kate Mckay, "What Drowing REALLY Looks Like," The Illustrated Art of Manliness, Sep 2020.

[4] Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai and Cewu Lu, "PMPE: Region Multi-Person Pose Estimation," Computer Vision and Pattern Recognition, Feb 2018.

[5] Chenge Yang, Zhicheng Yu and Feiyu Chen, "Human Pose Estimation Benchmarking and Action Recognition," Northwestern University, 2019.

[6] Rodman Louis,"Drown Detection Using Computer Vision," HKU Sport AI Lab, Nov 5 2019.